

ED 341 236

FL 019 921

AUTHOR Edwards, Jane A.
 TITLE Transcription and the New Functionalism: A Counterproposal to CHILDES' CHAT Conventions.
 INSTITUTION California Univ., Berkeley. Inst. of Cognitive Studies.
 SPONS AGENCY Alfred P. Sloan Foundation, New York, N.Y.; Max Planck Inst., Nijmegen (Netherlands).
 PUB DATE Mar 89
 NOTE 45p.
 PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
 DESCRIPTORS *Child Language; Cognitive Ability; Computational Linguistics; *Data Analysis; Databases; *Language Acquisition; *Language Research; Psycholinguistics; *Standards; Structural Analysis (Linguistics)
 IDENTIFIERS *Child Language Data Exchange System; *Transcription

ABSTRACT

The CHAT transcription standard of the Child Language Data Exchange System (CHILDES), the world's largest computer archive of child language data, is critiqued. CHAT is being considered as a possible standard for the field of child language as a whole. It is argued that CHAT is biased toward three views that are no longer current in the field: (1) toward written rather than spoken language; (2) toward a reductionist, compositional perspective on language and discourse; and (3) toward the view of child language as defective adult language rather than as systematic in its own right. A minimalist standard is proposed that is consistent with mainstream tenets, less theory-committed, and more readable than CHAT. It also can be accommodated by CHILDES' existing computer programs, and it enables exhaustive searches necessary for unbiased results. The minimalist proposal advocated is summarized in the appendix, which also contains tables on the following: CHAT's utterance delimiters; an example of a prosodically-based alternative to CHAT's unit-final delimiters; a comparison of Bloom's original format and CHILDES' version of the data in CHAT format; Bloom's data modified as necessary for line-oriented searching; CHAT's "special form" tags; and a sample from Clark's data on CHILDES in a non-CHAT minimalist format. (LB)

 * Reproductions supplied by EDRS are the best that can be made *
 * from the original document. *

ED341236

**TRANSCRIPTION AND THE NEW FUNCTIONALISM:
A COUNTERPROPOSAL TO CHILDES' CHAT CONVENTIONS**

By

Jane A. Edwards
*Cognitive Science Program
University of California at Berkeley*

**Cognitive Science Program
Institute of Cognitive Studies
University of California at Berkeley**

March 1989

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

- ☒ This document has been reproduced as
received from the person or organization
originating it.
☐ Minor changes have been made to improve
reproduction quality.

• Points of view or opinions stated in this docu-
ment do not necessarily represent official
OERI position or policy.

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

Edwards,
Jane

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

The production of the Berkeley Cognitive Science Report Series is sup-
ported by a grant in cognitive science to the University of California at
Berkeley from the Alfred P. Sloan Foundation.

FL019921

Abstract

The largest existing computer archive of child language data has a transcription method (CHAT) which—in the absence of explicit counterproposals—is being considered as a possible standard for the field of child language as a whole. In this paper, it is argued that CHAT is biased toward three views which are no longer current in that field, namely: toward written rather than spoken language; toward a reductionist, compositional perspective on language and discourse; and toward the view of child language as defective adult language rather than as systematic in its own right. CHAT's lack of systematic treatment of basic prosodics and conversation flow makes CHAT inappropriate for research concerning discourse, input effects, information flow, and language processing in real time. CHAT is problematic also for non-discourse research, due to technical properties—ad hoc categories, overanalyticity and non-uniform approach to regularization—which impede exhaustive computer search. In its place this paper proposes the minimalist standard widespread in the literature, supplemented by a uniform approach to regularization. This standard is consistent with mainstream tenets, less theory-committed and more readable than CHAT. It can also be accommodated by CHILDES' existing computer programs, and its minimalism and systematicity enable the exhaustive search necessary for unbiased results.

TRANSCRIPTION AND THE NEW FUNCTIONALISM: A COUNTERPROPOSAL TO CHILDES' CHAT CONVENTIONS¹

Jane A. Edwards

*Cognitive Sciences Program
University of California, Berkeley*

INTRODUCTION.

The Child Language Data Exchange System (CHILDES), headed by Brian MacWhinney (Carnegie-Mellon) and Catherine Snow (Harvard University) is the largest computer archive of child language data in the world. Since its choice of transcription standard will affect the research of generations of investigators, and since transcription is a heavily theory-laden enterprise, this is a choice of the utmost importance.

At the meetings in 1984 of the CHILDES Advisory Board in Concord, we² agreed on a set of criteria which a transcription standard would have to meet in order to serve the needs of the child language community at large, namely, that it would have to be:

1. a minimalist transcription system, involving only the most essential markings, in order to make it:
 - a. flexible for a range of theoretical perspectives and research questions,
 - b. easily readable, for visually scanning the data, and
 - c. minimally time-consuming to learn and to use.

as well as:

2. consistent with mainstream tenets of child language theory,
3. computationally tractable, and
4. expandable for user-specific needs.

In a number of ways outlined in this paper, CHAT has failed to meet these criteria. CHAT is a baroque system which engages the researcher in decisions which run counter to three of the main tenets of current child language theory, and which differs radically from minimalist methods emerging during the past decade as a tacit standard in the literature.³ CHAT's biases (including a bias toward written rather

¹ This work was supported in part by a grant to the University of California, Berkeley Cognitive Science Program, by the Alfred P. Sloan Foundation, and in part by the Max-Planck-Institut für Psycholinguistik, Nijmegen, The Netherlands.

² Elisabeth Bates, Ursula Bellugi, Lois Bloom, Melissa Bowerman, Robin Chapman, Eve Clark, Susan Ervin-Tripp, Paul Fletcher, Willem Levelt, Brian MacWhinney, Jon Miller, Ann Peters, Dan Slobin, Catherine Snow, and the author (see MacWhinney & Snow 1985: 276)

³ For example, Crystal (1975), Halliday (1975), Chafe (1980), Gumpers (1982), Ochs-Keenan (1983), Tedlock (1983), and several other American and British systems compared in Ehlich & Switalla (1978). Examples published after 1985 are cited later in this paper.

than spoken language) pervade the entire system, and cannot be eliminated by cutting CHAT back to only its utterance lines, since its biases are inherent in the transcription of individual words, division of interaction into units, and the (lack of) marking of prosodic information needed for discourse research.

CHAT's striking departures from the emerging standard in the literature are not motivated by computational considerations. In fact, several aspects of CHAT—overanalyticity, ad hoc categories, highly variable syntax and non-uniform approach to regularization—actually hinder reliable and efficient computer search, including the exhaustiveness necessary for unbiased results and generalizable conclusions. Another practical matter concerns the stability of the system in the future. CHAT's categories have already shifted several times, whereas the minimalist alternative, because it is more streamlined and less theory-committed, has changed relatively little during the same period of time, and will not become outdated as quickly.

Up to now, CHAT has been circulated informally, adopted by some researchers in the belief that CHAT format was required by CHILDES' programs. CHILDES' software can run on both CHAT and non-CHAT files, however (MacWhinney, personal correspondence, March 1988), and in fact one-third of CHILDES' enormous data base is not in CHAT format, including MacWhinney's own data in Hungarian. Clearly, the researcher has more choices available in transcription format, and it is important that the researcher know what those choices are and what each entails theoretically.

To the present time, CHAT has never been raised for public discussion, and has never been systematically compared with the minimalist standard based on intonation or tone units which pervades the literature. This paper is intended to address that need.

In what follows, I enumerate three main tenets of current child language theory, and show how CHAT is biased against them. Then, I outline the technical flaws in CHAT which prevent searches in CHILDES from being reliably exhaustive, as is necessary to ensure that the retrieved utterances are not mere a (potentially biased) sampling of relevant utterances. I show how the minimalist standard supplemented with a uniform approach to regularization satisfies both theoretical and computational requirements. In the process, I enumerate fundamental principles of systematic description which are satisfied by the alternative system and standard methodology in the field, but not by CHAT. The goal is to provide the reader with not only an awareness of sources of bias and how to avoid them, but also a positive alternative to CHAT, and an understanding of computational requirements.⁴ At the minimum, the user of CHILDES should be offered a choice between the CHAT system and a minimalist alternative of the type outlined below (see Appendix).

⁴ The earlier version of these comments, submitted to CHILDES in November 1987, resulted in minor changes in the May 1988 version of CHAT. The major issues remained unaddressed, showing that some of them may be overlooked even by experienced researchers. I hope that this paper will be a contribution to a continuing dialog on this subject.

TRANSCRIPTION AND THEORY.

CHAT's extreme departure from the Advisory Board 1984 guidelines and the emerging standard in the literature, may be partly due to underestimating the potential dangers of transcription bias for various types of research. Category-based or coding-based research is not much affected by transcription bias,⁵ but this is the exception rather than the rule. Many child language researchers depend on reading line-by-line through the transcript for insights, further hypotheses, and getting an intuitive feel for the nature and dynamics of the interaction. These researchers are highly vulnerable to inappropriate biases in a transcription system.

A transcript is an artefact—a distillation of the oral and gestural events of an interaction, expressed in a written and spatial medium, with properties of its own. The choices of which aspects to preserve, which categories to use, and how to visually format the data give rise to biases when reading through the transcript, which make some aspects of the interaction easy to examine, others more difficult or impossible. For example, relative visual prominence (upper/lower case letters, use of parentheses, etc.) affects perceived importance of different aspects of the interaction, and spatial proximity gives rise to the impression of interrelationship, whether intended or not. As noted by Ochs (1979), arranging successive child and adult turns one above the other gives rise to the expectation of mutual contingency, while arranging them in separate columns, with the child's utterances to the left, makes the child appear more dominant in the interaction (Ochs 1979). (For further discussions of introduced bias, see Tannen 1981; Brown & Yule 1983, p. 11; Preston 1985; Pye 1988; and Edwards *in press*).

Because bias cannot be avoided entirely, it is crucial that conventions be carefully selected so as to keep bias to a minimum and to ensure that those biases which remain are at least consistent with (i.e., does not contradict) what is known to be true about the type of interaction under study (see Ochs, 1979). All of transcription is a compromise, but some compromises are clearly better—i.e., less theory-committed or more compatible with existing theory—than others. Ignoring biases does not eliminate them, but only allows them to operate undetected.

Finally, there is the more basic methodological point of whether any discipline can afford to express its central data in a form which is biased against what it knows to be true, especially when an equally practical alternative exists which is consistent with its tenets.

In the past, changes in theory have been paralleled by changes in transcription practices. For example, in the 1960's, when the basic goal was to identify milestones in the development of an unfolding innate competence, the crucial data were the morphemes and words produced by the child, often without restarts,

⁵ Such research is, however, seriously harmed by unsystematicity and unreliability of category use, which impede the exhaustiveness of computer search necessary for unbiased research results. These difficulties in CHAT, arising partly from overanalyticity, are discussed below.

dysfluencies or prosodic information. Adult utterances were often edited for gist, and contextual information was limited or non-existent. In the 1970's, as emphasis shifted toward input and discourse context, older transcription methods became obsolete, because they lacked information which had suddenly become important or because they obscured dynamics in other ways (see Bloom forthcoming).

In the 1980's, there is an increasing focus on child language as spoken language, processed in real time, situated in a social milieu, partly shaped by the pragmatic functions it serves, and acquired in dynamic interaction with others (i.e., the "new functionalism"). As a result, naturalistic data predominate over experimental data, and computers have become essential for detecting regularities in the often enormous transcripts.

CHILDES, the largest existing computerized archive of child language data, has made important contributions in giving widespread, affordable access to computer search techniques adapted from various sources, as well as reading onto computer for the first time many of the historically most important data in child language. CHILDES' transcription method, CHAT, however, is problematic for both theory-related and technical reasons, as is discussed below.

CHAT's bias toward older theoretical perspectives, is no doubt due to CHILDES' trying to bring transcripts from several decades under a single common set of conventions. These transcripts could have been made suitable for computer processing with only minor adjustments. The difficulty with CHAT is that it sought to standardize conventions across three decades at a substantive level, and theory has changed since the 1960's. Transcription in the 1980's must be compatible with theory in the 1980's.⁶

CHILD LANGUAGE IN THE 1980'S.

Three tenets central to child language research in the 1980's, which need to be addressed in selecting a transcription standard, are that:

1. what is acquired is a *spoken* language, packaged prosodically, and processed in *real time*.
2. child language is highly contextualized, and is used intentionally to accomplish goals, and is acquired in the course of dynamic interaction with others.
3. the child's language is a system in its own right at each point in time rather than a deficient version of the adult language.

An adequate transcription standard needs to preserve not only the words and morphemes said by the participants, but also some systematic aspects of *how* they were said—basic prosodics and conversation flow—and must preserve this

⁶ This criticism pertains both to the application of CHAT to data in the 1980's and CHILDES' imposition of CHAT categories on data from earlier decades (e.g., the Bloom data and Brown data), that is, it applies to both directions in time. These are treated in turn in what follows.

information in ways which are easily readable and computationally tractable. The following discussion shows how these substantive requirements are satisfied by a minimalist alternative and how they are not satisfied by CHAT.

1. SPOKEN language, processed in REAL TIME.

Spoken language is organized into intonation units (or "tone units"), i.e., stretches of words which fall under a single intonational contour (Tedlock 1983; Chafe 1987, 1989) whose pitch trajectories (rising, falling, steady) are important for such things as information packaging, illocutionary force and turn taking (Eblich & Switalla 1976; Gumpers 1982; Jefferson 1984; Chafe 1987, 1989; DuBois, Cumming & Schuetze-Coburn 1988).⁷ Such contours have been found useful for a variety of languages (Chafe 1980), and are known to be salient to children as early as the one-word stage (e.g., Brown 1973; Galligan 1987). For example, children learning diverse languages tend to use rising intonation to mark yes-no questions long before using segmental means to do so, suggesting that "[I]ntonation is more easily grasped than rearrangements of order or the affixation of particles" (Brown 1973: 180).⁸

Local fluctuations in pitch or loudness in the adult language are informative with respect to word class (e.g., nouns versus prepositions), syntactic boundaries, and topicality (Lea, 1980; Waibel, 1987). There is evidence that they are also salient to children (Slobin & Welsh 1973), and might therefore serve as important cues in comprehension of unanalyzed wholes and analysis into parts (Peters 1983), regardless of whether or not they are used the same way as in the adult system (see Cutler & Sweeney 1987). For these reasons, it seems important to preserve basic prosodics in a field-wide transcription standard.

At a minimum, it seems necessary to segment the discourse into intonationally-defined (rather than syntactically-defined) units, to distinguish among several basic types of unit-final contours, to mark pauses, truncation of words and contours, overlaps (i.e., simultaneous speech by two or more speakers), and extra prominence of a word or syllable (see Appendix). Examples of work employing these conventions (and sometimes also fuller indications of sentence accent, and lexical

⁷ These references include work from linguistic, anthropological, and sociological approaches concerning adult discourse, both here and abroad. Despite differences in their specific research questions and explanatory frameworks, all three disciplines acknowledge the scientific primacy of prosodic and intonational information over higher-level coding (e.g., illocutionary force) in transcribing adult discourse. For child language, in which researchers must be especially cautious not to prejudge results, a focus on prosodics rather than more adult-centered syntactic or speech act categories is even more necessary.

⁸ Such categories can be detected reliably by speech recognition machines, even on the basis of only pitch contour information: "Based on their pitch contour, 77.8% of question and statement contours were detected accurately in 198 utterances. Only 3.3% errors are true classification errors and are not caused by segmentation or pitch tracker errors" (Waibel 1988:147). Discourse researchers sometimes titrate judgments of contour with normative expectations, allowing more variability in the acoustic signal than the machine might, but even at this, such decisions are less theory-committed than the judgments of illocutionary force, etc., required in CHAT.

stress) include the following: Leech and Svartvik (1975), Svartvik and Quirk (1980), Gumpers (1982; forthcoming), Ochs-Keenan (1983: 25), Jefferson (1984), Fletcher (1985: 14-20), Cruttenden (1986), Chafe (1987; 1989), DuBois, Cumming, and Schuetze-Coburn (1988). Many others could be cited.

CHAT, however, treats spoken language as written language. Its units more closely approximate sentences of written language⁹ than they do prosodically-defined units, and its seven unit delimiters (see Table 1) mark a mixture of (higher-order) functional, syntactic, and pragmatic distinctions rather than systematically marking basic prosodic contours.

Insert Table 1 about here

Five of them are treated as if they are conceptually mutually exclusive and exhaustive when they are not (in that only one is allowed to occur per utterance). Many of them are more interpretive (i.e., more theory-committed) and more specific than necessary and more committal than desirable in a field-wide standard. Prosodics per se are treated only very late in the manual (p. 18), and at that only as incidental characteristics rather than as highly perceptually-salient aspects, central to how spoken language is packaged into units in discourse, and processed in real time.

Rather than marking the distinction between rising and falling intonation, which is so customary in discourse research, CHAT's first two categories seem intended to mark illocutionary force, that is, a higher order interpretive inference concerning the speaker's intent in producing the utterance. Question mark stands roughly for *interrogative*:

- (1) rising intonation whether syntactically questions or not; or
- (2) syntactic questions spoken in non-rising intonation ("Where's the dog?").

Period stands for a mixed category best characterized as *declarative plus miscellaneous*:

- (1) "an unmarked (declarative) utterance," and
- (2) anything which does not fit the other six nonorthogonal categories.

In inferring how the utterance was intended (e.g., declarative? interrogative?) the researcher may resort to analogy with adult intonation-function correspondences, which are not necessarily the same as those in the child's system. Marking intonation contours instead, would preserve information known to be salient for both adults and children (even though specific function mappings differ—see Halliday 1975), without becoming committed to adult-oriented functional categories.

⁹ Despite the intent to be neutral in this regard—"The CHAT coding scheme does not attempt to tell the transcriber how to decide what should constitute an utterance" (CHAT manual, p. 11)—a definite bias toward writing-based sentence-like units is imposed by the nature and distribution of CHAT's utterance-final delimiters and other characteristics (see below).

Because CHAT's *period* category covers unrelated concepts (i.e., declarative plus miscellaneous), it is furthermore ambiguous in a transcript. Such categories obscure rather than clarify yet do so under a guise of precision which impedes later correction. More transparent would be to use a separate category for "miscellaneous."

"Steady" (or "non-falling/non-rising") intonation, often marked by discourse researchers with a comma, is marked in CHAT, but not as an utterance-final delimiter. In CHAT it is treated exclusively as an "utterance *internal* juncture." Yet comma intonation often occurs between intonation units, that is, in unit-final locations (e.g., when holding the floor for a lengthy narrative). Restricting commas to utterance-internal locations, and question marks, exclamation points, and periods to utterance-final locations is another manifestation of CHAT's bias toward written language. The distinction between utterance-internal and utterance-final delimiters is not necessary in an approach in which both the units and the unit-final delimiters are prosodically-based.

While not preserving basic prosodics, CHAT requires the researcher to distinguish five types of utterance incompleteness. The first two categories differ in terms of the presumed cause of the incompleteness: *interruption*, if due to factors external to the speaker; *trailing off*, if due to factors internal to the speaker. Unclear cases are marked as *trailing off*. Such judgments go beyond the data and may be unreliable with children, who may, for example, orient to different environmental noises than adults (cf. Scollon 1976: 30). In addition, similar to *period* above, *trailing off* is a mixed category and is therefore ambiguous in the data, potentially misleading and resistant to later refinement.

In addition to these two categories of incompleteness, if the same speaker resumes speaking after the incompleteness—in itself an interpretive judgment, especially in the absence of systematic attention to prosodic contour—there are three further categories to choose from: retracing with "correction," retracing without "correction," no retracing. This amounts to redundantly coding information which is already preserved in the flow of morphemes in the transcript,¹⁰ and doing so in unnecessarily theory-committal terminology. "Correction" implies an *error* which got fixed. This terminology is questionable when applied to adult productions, let alone to children, for whom the judgment will likely also involve appealing to an adult standard and an implicit deficit-oriented perspective on child language. Furthermore, although any of these three events could co-occur with either *interruption* or *trailing off* (as when the speaker resumes speaking after lapsing into silence or after a loud noise) CHAT would allow marking of only one of each pair of relevant categories (i.e., either *interruption* or *retracing with correction*).

All five of the categories for incompleteness are more specific than would be needed by most researchers, would require extra decision time, and tend to distract

¹⁰ CHILDES has recently announced availability of a computer program to add the second of these codes automatically.

from the more basic event they ought to signal, namely, simply "incomplete contour" or "truncation." It is only the latter which would be marked in the approach advocated here (see Appendix).

Emphatic stress, signalled by a complex of several acoustic properties, is defined in CHAT with reference to *imperative mood*, although in actual discourse, it can co-occur also with declaratives and interrogatives (e.g., "I won a million dollars!" "You want it done when?!"). Given CHAT's constraint of one "utterance-final delimiter" per unit, if **emphatic stress** occurs in a declarative (**period**) or an Interrogative (**question mark**) utterance, presumably only the more "marked" (in the linguistic sense) alternative—in that case, **emphatic stress**—is preserved in the transcript, and similarly for combinations of interruption with declarative, or **trailing off** with declarative. This means that unit-final delimiters are assigned according to some hierarchy. But this introduces further ambiguities in that the overshadowed category is not transcribed in a systematic way and could not be retrieved reliably in computer search. If one is going to use non-orthogonal categories, then it is more sensible conceptually and safer methodologically to use "multiple marking." That is, if more than one category applies, mark more than one.

Latching and completion pertain to **discourse continuity**, and presuppose assumptions which may be overly committal for a field-wide standard for child language research. **Latching** presupposes a tacit norm of "customary" pause length between utterances. This norm varies widely even across different conversational styles in the same language for adults (Tannen 1984). It is not known how it varies during child language acquisition. **Completion** is biased toward the written language, and toward sentence-like units. Spoken language is full of structures which would be called "fragments" in written language, so it would often be difficult, even with systematic marking of prosodic contours, to distinguish between completion of an earlier turn and initiation of a new turn concerning an already-shared topic.

In summary, CHAT seems to treat child language as being similar to written language. Its units are like sentences, and it lacks systematic preservation of prosodic information necessary for viewing child language as a spoken language, which is processed in real time. Rather than marking prosodic contours, CHAT's unit delimiters involve higher-order coding decisions many of which would not be of use to individual researchers because they are too detailed or unsystematic. CHAT's utterance-final delimiters are treated as if they were orthogonal when they are not. They are also not exhaustive, except in the artificial sense of having some conceptually motivated categories serve also as "miscellaneous" categories (**period** and **trailing off**). Such categories are conceptually impure, and ambiguous in transcriptions, while at the same time giving the impression of precision, which hinders later refinement, even with access to the tape recordings—something CHILDES considers usually unnecessary after the transcript is "complete" (see below). In place of CHAT's delimiters, more prosodically based distinctions are proposed, including

those shown in Table 2.

Insert Table 2 about here

These same distinctions are found in Chafe (1980), Jefferson (1984), DuBois, Cumming and Schuetze-Coburn (1988), Gumpers (1982), and Tannen (1984) (though some are marked by different arbitrary symbols).

2. Conversational Flow and a Non-Compositional view of Discourse and Language.

Information concerning temporal interleaving and pacing of verbal and nonverbal aspects (including prosodic, contextual, gestural and other aspects) is necessary if we take seriously the view that the child's language is highly contextualised, used for communicative purposes, and acquired in the course of dynamic interaction with others. To preserve the perceptual unity of the communicative event, and to give a clear sense of how that particular event fits into the larger conversational flow, discourse researchers tend to transcribe the event in a maximally compact manner, keeping verbal and non-verbal¹¹ aspects together on the same line, distinguishing them by means of parentheses, type font or other devices. For example:

*SAR: she's goin' (=going) come back. ouch! ((yelling))

In contrast, CHAT encourages spatially isolating the verbal from non-verbal and then further subdividing non-verbal into up to twenty separate lines called "dependent tiers," each of them preceded by a theory-committal label (pp. 61-62). For example, the CHAT version (p. 23), of the above example has an added %par, or "paralinguistic" tier:

*SAR: she's goin(g) come back. ouch!

%par: <2u> yelling

In the CHAT version, the temporal relationships are not represented iconically as they are in the more discourse-oriented version. Rather, they must be computed by the reader from a numerical index ("<2u>" for "second utterance").

In addition, the spatial isolation of parts implies a compositional view of discourse, namely that the different aspects are functionally dissociable, and that the communicative event is the sum of its parts. This conflicts with current mainstream perspectives, namely, that the different aspects may often serve overlapping functions, and that what is important is how they interrelate as part of a *unitary*

¹¹ In what follows, I use the term "non-verbal" to mean "anything other than verbal," and "nonverbal" to refer to the subset of "non-verbal" pertaining to motion or position of the body or body parts.

communicative event. They interact in the interpretation and contextualization of that event in ways which affect the course of later utterances in the discourse (Erickson & Shultz 1982; Gumpers 1982).

The only time that separate tiers are needed for practical reasons is when descriptive comments or codes are either so lengthy that they would disrupt readability, or when they pertain to more than one communicative event (e.g., shifts in situational context, on-going activity, etc.). This is not the case for the majority of situations treated by CHAT.

Apart from imposing a compositional view, CHAT's use of "dependent tiers" subordinates non-verbal to verbal, making non-verbal seem incidental, rather than communicative in its own right. Thus, the only time that there is any indication on the utterance line that a gesture occurred is when the gesture occurred unaccompanied by words, in which case it is marked by "0." on the utterance line (presumably for "zero words"). For example (p. 14):

*FAT: where's your doll?
 *DAV: 0.
 %act: [runs into bedroom]

where the contents of the %act, or "activity," tier supposedly replaced an utterance by DAV.

All other times, the gesture or activity simply appears on a later tier without any indication on the utterance line. For example (p. 22):

ANN: Mommy I want to go <ger> [].
 %err: ger /gEr/ -> there /dEr/ \$=/ {g}o/ \$SUB \$CON I
 %par: whining
 %int: <1> rising
 %exp: <6> to the store
 %gpx: points to the door

where %gpx stands for "gesture, proxemics."

But, Ochs (1979) and others have noted that non-verbal aspects of the communicative gestalt may be as important as verbal or even more important for certain purposes, and that it is misleading to subordinate them arbitrarily to the verbal stream. She avoided even setting them off in parentheses on the utterance line; the CHAT conventions imply much more extreme subordination.

The extreme analyticity found in CHAT is also not dictated by technical considerations relating to the use of computers—in fact it actually hinders efficient and exhaustive computer search. A researcher concerned with deictic expressions would be interested in the gesture on the %gpx in the preceding example since it partly disambiguates the *there* on the %err (for "error"¹²) tier, which is the regularized version of "<ger>" from the utterance line. In searching through a CHAT

¹² The adult-centeredness of the %err tier is discussed below.

transcript, the researcher would need to retrieve all %gpx tiers whenever a deictic expression occurred on either the utterance or %err tiers. Even this would not be enough to ensure exhaustive search, since CHAT recommends not one but three separate tiers for nonverbal events:

- %gpx for gesture, proxemics,
- %act for actions (e.g., "runs to toy box"), and
- %fac for facial actions (no examples given).

The researcher would have to retrieve at least five different types of tiers to study deixis, even though many individual instances might turn out to be irrelevant to deixis.

Even if they did not hinder readability and uniform computer search, the large number of labelled dependent tiers (gestures, action, facial actions, etc.) would be undesirable because they are difficult to distinguish in practice, and because they commit the researcher to not only a compositional view of discourse but actually a particular theory-laden set of categories which may be largely irrelevant to his or her own research goals or beliefs.

In contrast, consider the simpler type of representation advocated here:

ANN: ((whining)) Mommy? I want to go </gEr/> (=there). ((points to the door))
 - <*"there"="to the store"> \$=/ {g}o/ \$SUB \$CON

The top line serves as a self-standing record of basic verbal and nonverbal aspects of the communicative event. The second line contains additional clarification and coding. This greatly lessens subordination of nonverbal to verbal, and is more compact, facilitating perception of discourse flow. Its compactness derives partly from the use of the ? for rising intonation on the utterance line rather than on a later tier subordinated to it. This single symbol placed on the utterance line, where it is more relevant, eliminates the need for: an extra tier, an index ("<1>" for first word), and an explicit descriptor ("rising"). Critical to note is also the absence of a theory-committal label for the second tier.

Sometimes additional information is necessary, such as the specific stretch of words in an utterance to which a comment applies. But even here, it is not necessary to use an independent tier, as in the following example, where the child giggles through the third to fifth word of her utterance:

*ANN: that clown was very funny . ((giggling(3-5)))

CHAT hinders perception of conversation flow not only through a proliferation of tiers, but also through the use of extra, and uninformative symbols on the same line. Consider, for example, CHAT's treatment of overlaps, that is, simultaneous speech by two or more interlocutors (from p. 20):

*MOT: no [#] Sarah [#] you have to <stop doing that!> [overlap>]

*FAT: <oh my gosh!> [<overlap>]

*SAR: <Mommy I don't like this.> [<overlap>] it is nasty.

In this example, paired angle brackets indicate the part of each speaker's utterance that was involved in the overlap, and the differential bracketting around overlap serves to tie them together as part of the same overlap event.¹³ The more discourse-oriented version accomplishes the same purpose but in a much more readable way:

*MOT: no ... Sarah ... you have to [stop doing that!]

*FAT: [oh my gosh!]

*SAR: [Mommy I don't like this.] it is nasty.

Alternatively, one could forego the vertical alignment, and bind the three together with a numerical index:

*MOT: no ... Sarah ... you have to [stop doing that! 1]

*FAT: [oh my gosh! 1]

*SAR: [Mommy I don't like this. 1] it is nasty.

Either of these is more readable than the CHAT example, and a clearer representation of conversational flow.

CHAT authors are aware of problems in readability of CHAT for discourse purposes, but propose to solve it by adding yet another tier (the %flo tier), stripped of all markings indicating how the utterance was said (pp. 22-23):

%flo simplified "flowing original" rendition of the main line

This line is designed to allow for a level of transcription that is as free as possible of transcription conventions and which forces the transcriber to make the minimal number of transcription decisions. Here is an example of the %flo line:

CHI: <I do-'nt> [/]/ I do-'nt wan(t) (t)a look in <a> [] <badroom>
[*] or Bill-'s room.

%flo: I don't I don't want ta look in a badroom or Bill's room.

%err: a /A/ -> the /DA/ | badroom /baedrUm/ -> bathroom /baeTrUm/

Though intended explicitly to meet the needs of conversation analysts, this is not what CA or other discourse researchers mean by "flow." They systematically mark prosodics, incompleteness, turn-taking, pauses, etc. and these things often figure

¹³ The differential bracketting serves no substantive purpose, such as identifying some speakers as being more prone to intrude on others, because the decision of which person's turn to transcribe as the first in the overlapped sequence—e.g., FAT or SAR—often depends on irrelevant and incidental factors (e.g., the relative length of the overlapping utterances, and whether or not they continue through and beyond the overlap). I am indebted to Deborah Tannen for this observation within the context of adult discourse transcription.

prominently in their analyses. In the absence of such marking the researcher is forced to impose speculations on the data in order to resolve frequent ambiguities. The %flo tier would not help but only further impede perception of conversational flow, by introducing needless ambiguities and adding yet another tier to be scanned. In contrast to the CHAT approach, the minimalist alternative involves capturing the utterance itself in a maximally readable way, minimising the number of separate tiers, and avoiding theory-committal labelling wherever possible.

Another aspect of CHAT which impedes gaining a sense of discourse flow is that contextual comments are routinely attached to and placed after a single turn or utterance rather than being set off from and placed before a turn or utterance, as is customary in discourse research and is more appropriate for capturing the context-sensitivity of utterances. That is, contextual information shapes and often precedes utterances. It is confusing to read transcripts in which this ordering (and seeming direction of influence) is reversed. Table 3 contains both the original and the CHAT version of Lois Bloom's (1975) data for comparison.

Insert Table 3 about here

What is described in Bloom's opening lines: (*M and A sitting on chair; A wearing half-zippered jacket; fingers in her mouth*) is not a set of features of the Mother's opening turn, but rather the context in which the discourse as a whole begins. Its scope of influence may well have been more than just the first turn in the discourse, especially since it includes background information concerning not just one but both speakers. Placing situational comments before and independent of verbal events allows the possibility of influence beyond the immediately subsequent turn without being committal concerning the exact nature and duration of their influence, which we often cannot know, but which is presupposed in the CHAT convention to be the immediately adjacent turn.

Bloom's original version makes it easy to see structure at the gestural level independently of structure at the verbal level. In the Bloom version, for example, it is easy to infer that the Mother's pointing to the monitor caused Allison's looking at the monitor—a sensible inference considering the nature of the nearby verbal events. In CHAT version, in contrast, this possibility is hidden in a list of unstructured features, split across several tiers, and subordinated to the spoken event.

This short example also shows extreme unreliability in distinguishing among CHAT's theory-laden separate tiers. "Fingers in her mouth" is treated initially as %sit, then as %gpx and then as %act. "Looking at the monitor" is treated as %gpx while "as she watches the monitor" is treated as %act. But even if such distinctions could be made reliably, they are of little practical value, since any researcher interested in context would want to read all of these lines, rather than selectively suppressing any of them. The theory-laden decisions, therefore, take time and distract the researcher, without providing anything of benefit in return.

Finally, in the CHAT comments, personal initials have been edited out, making them locally ambiguous. The only way the reader knows who is being described is by backtracking to the most recent utterance line. This takes time and greatly complicates the task of gaining a sense of the flow of the conversation. Compare, for example, the CHAT %gpx tier: "<aft> points to monitor" with Bloom's version: "M points to monitor," in which the referent is explicitly stated.

For computer analysis of Bloom's data, all that was strictly necessary was that the three types of events comprising the interaction: comments, M's utterances and A's utterances, be separately addressable or suppressible (plus regularization treated in the next section). They could have been left in columnar format, or reformatted strictly mechanically¹⁴ (i.e., with no added markings or distinctions) so that the different events appeared on separate tiers of equal dominance, preceded by a label (i.e., a line-initial identifier which indicates which of the three it is) (as I have done, as an example, in Table 4).

Insert Table 4 about here

The addition of CHAT transcription distinctions—especially in the absence of access to the original tapes—goes far beyond what was necessary for computational analysis of the data, and increases the margin of error of all the categories used. Imposing CHAT categories on the basis of superficial similarity is risky, and the resultant uniformity makes researchers less aware of actual differences between data sets which may actually be important for interpretation.

3. Child-Centered rather than Adult-Centered Description, and a Uniform approach to Regularization.

In order to enable efficient readability and exhaustive computer search, it is essential to associate a predictable form (e.g., standard spelling for a lexical item) with each aspect of the interaction which the researcher might wish to retrieve, a process called "normalization" or "regularization."¹⁵ The categories used must be relevant to the child's own system.

Since the 1960's, the child's system has increasingly been described as a system in its own right, not unlike a language examined in linguistic field work (or "language variety"). The field researcher is careful not to impose inappropriate categories from his own language on another language, and avoids comparing isolated elements from the two systems (i.e., languages) without careful consideration of the functions each serves within its respective system. A similarly cautious methodology is pervading child language research and must be reflected in a field-

¹⁴ The typographical error of "ALL:" in place of "ALL:" suggests this was done by hand rather than by computer program, which may mean the possibility of additional extraneous error.

¹⁵ Technical details are discussed below. The focus in this section is on the substantive issues involved in choosing appropriate categories.

wide transcription standard, but is not present in CHAT. CHAT is shown below to employ adult-centered categories, to encourage comparisons of isolated elements of the child and adult systems, and to employ a partly conflicting set of regularization strategies which hinder readability and exhaustive computer search with no commensurate substantive or analytic gain. In this section we examine the substantive biases involved in CHAT's approach to regularization. Technical details, including computational details are discussed below in a separate section.

Concerning regularization, CHAT authors recommend¹⁶ that (p. 10):

The speech on the main line must be *regularized to standard [words and] morphemes in the adult language*. That is, the goal is to enter on this line the most likely target [words and] morphemes in the adult language. It is important to emphasize that this regularization is *only on the morphemic level and not on the syntactic level*.¹⁷ There is no need to correct syntactic errors on the main line. Syntactic errors are indicated separately on coding lines. There is a mechanism for noting omissions on the main line. However, not all researchers will want to code all omissions. In fact, some researchers may feel that they cannot decide what kinds of words have been omitted and therefore prefer not to code omissions at all. [...] If the word has an error in it, the target form is given on the main line and marked with the error symbol [*]. (*italics theirs; emphasis mine*)

Assimilating the child's linguistic forms to adult words is a common practice and has face validity, since children appear to communicate with chunks as large as words or even larger. It is also a sensible choice since most researchers are not interested in phonological variation except, for example, where it may reflect points of reorganization in the child's system (for example, when an American child uses *cannot* rather than *can't*, as if to make the negative element more regular and perceptually prominent).

Assimilating the child's linguistic forms to adult morphemes, however, is highly controversial and theory-committed. It goes beyond the necessary level of standardization, and imposes adult categories on the child's speech which may well not yet be relevant in the child's own system, especially at early levels (i.e., parts of unanalyzed wholes), or may well have extensions different from those in the adult's system (by virtue of contrasting with a different set of items in the child's system). CHAT's recommendations impose a definite adult-centered bias.

Adult-centeredness is also shown (in the above quote) in CHAT's referring to any departures from adult syntax as errors, and omissions, which the researcher

¹⁶ The bracketted "words and" were clearly intended here, based on preceding context, even though omitted twice in this passage.

¹⁷ CHAT's caveat against regularizing on the syntactic level is violated by treatment of "non-completed" elements, discussed below.

correct[s] rather than viewing them as legitimate forms within the child's own language variety and coding the functions they serve in the context of the child's own system. It is also shown more subtly in choice of phrasing: most likely (rather than "nearest" or "most similar") and decide what words have been omitted, which presupposes some have indeed been omitted. To avoid excessive influence from one's own language, it is best at pre-theoretical stages to preserve both the surface form and the coded category.

CHAT uses six different partly conflicting, partly overlapping conventions for regularization:

- (1) Adult morphemes.
- (2) Phonological and syntactic "errors."
- (3) "Non-completed" words.
- (4) "Omitted" words.
- (5) Language mixing.
- (6) "Special form markers."

These are examined in turn.

(1) **Adult Morphemes.** CHAT (p. 16) advises a morphologist to replace a child's *he's* with either *he-has* or *he-is* on the main line, depending on which of these would have fit the context according to the idealized adult standard. Similarly, *let's* is to be replaced with *let-us*. This practice involves capturing the child's utterances in the form of adult categories which may not be relevant to the child until he actually achieves literacy. Preferable would be to preserve the child's actual utterance along with the relevant morpheme codes from the adult standard, enclosed in parentheses to enable optional suppression:

let's (= 'us)
and
he's (= 'has) got a fish.

The equals sign marks the correspondence between spoken and non-spoken (coded morpheme) items in a visually prominent way. The apostrophe before the coded item (e.g., 'us) eases computer search for morphemes.¹⁸

Rather than placing the standard version immediately beside the nonstandard variant, it would be possible to ensure the transcript was exhaustively searchable by providing a conversion table of nonstandard-to-standard forms either at the top of

¹⁸ CHILDES' data files (as of the version distributed January 1989), contain a number of omission of either the dash or the apostrophe in their conventions or both, making these parts of CHAT codes not a reliable basis for search in their data. As just one example, the double possessive *mine's* is transcribed as *mine-'s@n* in CHILDES' version of Brown's Adam data, but as *mine's@n* in Brown's Eve data. Or, compare the following examples of inconsistencies from CHILDES data:

"he's learning how to say his aOl bOl c@l-s"
"go write some aOl bOl c@ls."
"draw aOl bOl c@l's."

the transcript file or in a separate file, especially if specialized software were made available which uses that table to expand the search command or automatically generate an "include" file (using CHILDES' programs), so that each of the variants is specified for search in a single run with a minimum of effort by the researcher (e.g., *till*, *til*, (*un*)*til*, *until*). Since CHILDES and many facilities lack the software to exploit a lexicon entry in this way, discussion below favors placing the standard version immediately beside the nonstandard variant, since this enables direct retrieval of all nonstandard and standard instances of a form, based on the presence of the standard form on the utterance line. Adjacent regularization also serves to disambiguate items which might otherwise be unrecognizable to the reader,¹⁹ without requiring the reader to refer elsewhere for clarification, and it makes it easier to detect nonstandard items as yet unregularized due to oversight. If the conversion table approach is used instead, occasional adjacent regularizations may be useful to clarify locally ambiguous items.

Except in the rare event that the child uses the same surface form for two very different lexical items, either type of regularization (adjacent or conversion table) can be converted into the other, so they are treated as functionally equivalent in what follows—i.e., as a single method with two implementational variants. Both require that regularization be exhaustive, that is, that for every different type of nonstandard variant, a standard form be assigned as a search key for its retrieval. For ease of exposition, I will compare CHAT with the adjacent regularization approach in what follows, but the elements placed in parentheses could just as easily be entries in a conversion table.

The method proposed here (implemented either by adjacent regularization or by means of a conversion table) can be extended with only minor modification to cover all six of the regularization strategies proposed by CHAT, and thereby constitutes a uniform solution to the problem of regularization. Such uniformity is desirable for reasons of greater predictability and therefore better readability and computational efficiency, in contrast to CHAT's more variable, and more piecemeal approach to regularization.

(2) "Errors." This CHAT convention involves placing on the main line "the form that was actually produced," enclosed in angle brackets, followed by [*] to identify it as an "error," and then clarifying it on a separate tier (%err). This method was used for *a* and *badroom* in the following example (p. 23):

CHI: <I do-'nt> [//] I do-'nt wan(t) (t)a look in <a> [] <badroom>
[*] or Bill-'s room.

%flo: I don't I don't want ta look in a badroom or Bill's room.

%err: a /A/ -> the /DA/ | badroom /baedrUm/ -> bathroom /baeTrUm/

The adult-centeredness of referring to departures from the adult system as "errors" has already been noted. A further problem is the complexity of the representation,

¹⁹ This is a common problem with eye-dialect transcripts, as discussed below.

and the presence of more detail than most researchers would need. The alternative proposed here would be to transcribe the end of that utterance instead as follows:

"look in a* (=the) badroom (=bathroom)"

The advantage of this method is that it preserves both the child's actual utterance and the standard form needed for exhaustive search, and does so with good visual separation between what was and was not actually said by the child. In addition, placing an asterisk on lexically or syntactically nonstandard forms (but not on purely phonological variants) enables efficient retrieval of informative departures from the adult system. Phonological variants could be retrieved without additional marking (along with other regularized but unasterisked forms).

(3) "Non-completed" words. This method is reflected in the (t) in *wan(t)* and in (t)a in the above example, which were not actually produced by the speaker:

"When a word is incomplete, but the intended meaning seems clear, insert the missing material within parentheses. This is also useful for indicating the omission of segments. It is also helpful for noting commonly used shortened words, such as (un)til and (be)cause" (p. 15). (italics theirs)

"Incomplete" is a judgment based on adult-centered perspective as well as written-language biased. The adult forms are overlaid on top of the child forms with such minimal visual separation that the line must be read very slowly in order to avoid attributing more adult-like forms to the child than he or she actually uses. Consider the following example (p. 15):

*MAR: I been sit(ting) all day.

If the form *sit* is a legitimate form within the child's own system, it seems it should be preserved clearly as such. This CHAT example appears to violate CHAT's own prohibition against regularizing at the syntactic level. If a regularization is to be provided in these cases, there may also be several adult forms which might have applied, depending on which semantic aspects were most salient for this child—something the adult researcher can only guess. For example, this child may have viewed it as an endstate (*sat*) rather than an ongoing activity (*been sitting*). The extra visual separation would lessen the perceptual bias toward one or another adult form.

The judgment in CHAT that something is "non-complete" is biased not only toward adult language, but actually toward adult *written* language. The form captured as *wan(t)* (t)a above is commonly *pronounced* by adults without the t's. It is so common a form in *spoken* language that it is increasingly transcribed as *wanna*. Such forms as *ta* or *tuh* for *to* are instances of what has been called "eye-dialect," or "allegro style" transcription, that is, the use of distorted English to preserve the flavor of the dialect (e.g., *wuz* in place of *was*, and *eye* in place of *I*) or some non-standard pronunciation. (see Edwards in press; Levinson 1983, p. 296). Such forms (e.g., *badroom* in place of *bathroom*) impede exhaustive search unless "regularized" (see below). Without regularization, a researcher performing a computer search for

infinitival (or *to*-) complements would miss the above example unless he or she knew in advance that the ("eye-dialect") variant, *ta*,²⁰ existed in the data, and remembered to specify it in the search.

The approach recommended here would use an extension of the method used for the morpheme example rather than totally different regularization conventions for the two situations:

I don't wanna (=want to) look in a* (=the) badroom (=bathroom).

namely preserving the child's spoken form together with the adult standard enclosed in parentheses to enable optional suppression. This method preserves readability, ensures exhaustive computer search, and is much less adult-centered than CHAT's third regularization convention.

(4) "Omitted" words. Under CHAT conventions, if a word was "omitted" by the child, the researcher is supposed to guess at what it *might have been*, and type it on the main line, prepended with a zero (p. 15):

*EVE: I want 0to go.

This convention, like two of the three which preceded it blurs the distinction between what was or was not said by the child. In quickly scanning the text, it would be easy to overlook the fact that the form did not actually occur in the utterance, something which would hinder particular types of insights into larger structural characteristics of the child's system. The following alternative is less biasing while still maintaining readability and enabling uniform search:

*EVE: I want ^ (=to) go.

(5) Language mixing. In CHAT, the researcher would need an extra tier in cases of language mixing (p. 24):

*HIL: Spiel house, knock down.

%ger: <1> game -> Spiel

as well as a positional index ("*<1>*" for "first word"), and an arrow. In contrast, the regularization method proposed above produces a one-line alternative:

*HIL: Spiel (g=game) house, knock down.

In this example *g* indicates the item was in German in an otherwise English-language utterance. This is more readable, and more efficient for computer retrieval than the CHAT version.

(6) Special form markers. For such things as 'neologism', 'child-invented form', 'babbling', and various 'shortened' forms ('*gain* for *again*'), the CHAT user is supposed to append a category code (e.g., @n, @c, etc.) to the end of the relevant spoken item and to list the item in a separate lexicon file, together with clarification

²⁰ This violates CHAT's principle that child forms be standardized to the "most likely" standard adult form.

i.e., glosses or standard forms, as a help to researchers in achieving exhaustiveness of the search.

As with other sets of CHAT categories, CHAT's Special Form categories are not mutually exclusive. Yet, only one tag is allowed per item even when more than one applies. This opens the door to inconsistencies in the way the tags are used in actual data.

Table 5 contains the full set of CHAT's tags found in the January 1989 version of the CHILDES data base. The examples were compiled from from CHILDES' "lexicon" files, and CHILDES' versions of Roger Brown's data and Brian MacWhinney's "boys" data, which can be seen as paragon cases, since tags were added to them by the same researchers who created CHAT and the Special Form categories.

Insert Table 5 about here

In the table, one finds *squished*, *squashed*, *smushed*, and *squoosh* tagged variously as 'phonological consistent form', 'child-invented form', or 'extended English' (yet for some reason not as 'onomatopoeia'), while *gooshy* is tagged as 'motherese'. Forms such as *teared* and *sitted* get various tagged as 'child-invented', 'neologism' or 'extended English'. Such inconsistencies play havoc with computer retrieval, and certainly mean that the tags cannot themselves serve as the basis for reliable search. They seem therefore to impair readability without providing any reliable benefits in return.

The 15 categories marked by tags in CHAT could actually be reduced to seven categories, all of which either are already handled by regularization conventions discussed in subsections 1 through 5 above, or else do not really warrant explicit tags, due to being too specific for inclusion in a field-wide standard, or due to being more readably handled by other means which are equally effective computationally.

The first class concerns spelling aloud. For this purpose, CHAT adds a @l, but far more readable and equally easily retrievable²¹ is the minimalist alternative of using a capital letter for each letter recited:

CHAT (from the Brown data):

*MOT: the s@l [#] t@l [#] o@l [#] v@l [#] e@l [#] so she always says

Minimalist alternative:

*MOT: the S ... T ... O ... V ... E ... so she always says

The second class concerns qualifiers, marked with @q in CHAT (e.g., *yep*, *haha*, *uhuh*), to which CHAT authors add descriptors, serving the function of

²¹ The only time ambiguity arises is with *I*, which is almost always clear from context. For the rare ambiguity, clarification is added in double parentheses.

regularization, to allow the researcher to retrieve all items with the same semantic or functional value in discourse with a single search specification. Most researchers, however, would never have the need to retrieve all tokens of the semantic item *yes* (e.g., *yep, ya, yeah*), nor all tokens of laughing (e.g., *haha, hahaha, hehe, hoho*), nor of exclaiming (e.g., *hey, ha, aw, ah, woooo, oo, ooh, wow, ahh, eesh, hm*). But if they did, this could be achieved with better readability by simply listing them exhaustively in the lexicon file associated with the appropriate descriptive gloss, that is, without a special form tag on the utterance line. Descriptors also almost never needed for readability, in transcripts which mark basic prosodic contours and preserve conversational flow in a systematic and readable manner. Compare, for example, *huh?* with rising intonation versus *huh.* with falling intonation. For this reason, there is no need for tagging of these in the minimalist approach. They are also left unglossed except for those rare instances which are ambiguous despite prosodics and discourse flow, in which case a comment is added in double parentheses.

A third class is nonlexicalised sounds, tagged in CHAT as @h, @b, @o, @p, and @w (i.e., hesitation, babbling, onomatopoeia, phonological consistent form, and word play). Tagging these forms constitutes a level of coding not needed for most research purposes and is therefore too specific for a field-wide standard, since it would increase decision time at data entry unnecessarily. They may also lessen readability, especially the tags involving vowels, which on quick scanning of the transcript, sometimes appear to be a part of the actual item spoken.

The fourth class is lexical items which are common in spoken language but inadequately captured with standard written orthography, e.g., *gonna*, or *Grampa*, tagged in CHAT by @f, @m, @e, and @a. A minimalist approach would not use special purpose tags here either, so long as each nonstandard item is associated with a standard form to enable exhaustive search, as discussed under subsection 1 above.

In the fifth class are lexical items which deviate from spoken adult norms, and which may illuminate structure and points of change in the child's system (e.g., *breaked, ated*), marked by CHAT's @n, @d, and occasionally @c. These are discussed under subsection 2 above. Appending an asterisk at the end of such items and associating a standard form with it either adjacently or in a separate lexicon file is adequate for purposes of uniform retrieval, and preferable for readability and expediency in data entry and search (due to lesser chance of irrelevant variability).

The sixth class, marked occasionally with @c in CHAT, involves novel lexicalizations, i.e., when the child uses a known adult form with some other meaning (e.g., novel causatives, lexical gaps, novel instrumentals). These require a level of effort and background knowledge such that marking them may not be possible for all researchers, and they should therefore be left optional. Researchers who do wish to mark them, however, could do so less intrusively by appending !—rather than the more distracting @c—with clarification given in double parentheses.

The seventh class concerns second language forms, @s, which are already covered under subsection 6 above.

I mentioned briefly the problem of variability. This is a serious bane to exhaustive search. As a transcript becomes more complex visually, there is often a danger that unmotivated variability increases because inconsistencies get swamped in the enormous mass of information the eye must process.

This "information overload" effect, is suggested by such things as the variability in the syntax associated with these tags (e.g., for plural forms: *deer@n-s* vs. *reindeers@n*; for possessives: *doggie's@m* vs. *kittie@m's*; for added glosses: *choochoo@m=train*, *mhm@q=yes*, *ya/yes@q*, *(un)til@a*), and also in the proliferation of synonymous descriptors (e.g., in CHILDES' version of MacWhinney's "boys" data, *yick*, *yuk*, *yucky* and *yukky* are glossed variously as *discust*, *discust*, *disgust*, and *bad*; *mm* is glossed variously as *delicious*, *tasty*, and *yum*; and the descriptor *yum* in turn gets glossed as *delicious*; the data include also *thinking* and *thoughtful*, *laugh* and *laughing*, *what* and *question*, *true* and *affirm*). The examples cited are also not isolated instances. Many others could have been cited. These types of variability pose the same hindrance to exhaustive search as nonstandard lexical items in the transcribed utterances. For either of them the researcher cannot be sure the search will even roughly approximate exhaustiveness, regardless of the computer programs available. This is another reason for preferring a more streamlined, and systematic transcription format.

Finally, the conversion table or "lexicon file" approach to regularization only works if: (a) every nonstandard variants in the data is included in the lexicon file, and (b) each of them is associated with a standard form the researcher might use in an attempt to locate them (e.g., *gonna*, *gon'* = *going to*). CHILDES' "lexicon" files at present are incomplete on both counts, which restricts their effectiveness in helping to achieve exhaustiveness of the search. The researcher may overlook relevant cases in visually scanning the incomplete list of nonstandard forms taken out of context, and attempting to guess which of them may have been child approximations to the standard form of interest in the search. Rather than resolving the regularization problem, it merely displaces it.

Summary concerning regularization. CHAT's six regularization methods are either adult-centered (four of them), or impede perception of conversation flow and computer search (two of them). One of them (morpheme marking) replaces what was said with adult-centered categories which are known not to apply uniformly throughout acquisition. Three of them violate CHAT's own principle against putting things on the main line which were not actually said. These three involve supplementing child forms with adult forms on the main line with such minimal marking that it is impossible to read the record quickly without having the impression that the child has used many more adult forms than was in fact the case. Some of this supplementation involves "correction" at the syntactic rather than merely morphological level—which is also inconsistent with CHAT's own dicta. One of them (special form tagging) does not solve the problem, but merely displaces it.

adding at the same time an extra level of theory-committal tagging. Finally, CHAT allows the occurrence of some nonstandard forms which never get regularized (e.g., the use of *ta* or *tuh* in place of *to*, which would be overlooked in a search for infinitival complements unless the researcher knew in advance of its existence in the data set and remembered to specify it along with other possible variants).

CHAT's use of six regularization methods where one would do increases variability in the transcript which slows reading and can cause the researcher to overlook relevant data, with the result that conclusions may be based on a biased sampling of utterances and may not generalize.

In place of CHAT's six methods, a single method is proposed which, with only minor modification, applies to all six contexts. It involves preserving what the child actually said in a systematic way (i.e., other than eye-dialect), and associating with it with the adult standard set off in a visually distinct manner (and optionally suppressible) in parentheses preceded by =, and (where needed) a one- or two-letter code (e.g., apostrophe for "morpheme," *g* for "German"). This reduces the danger of overlooking relevant data since the standard form will always be locatable on the main line, rather than sometimes being found there and sometimes on dependent tiers, expressed in a range of different notations.

The use of an exhaustive conversion table, associating a standard form with each nonstandard form in the data, is an implementational variant. Either way, *teared* would be marked by an asterisk and a standard version would be provided to help in exhaustive search. Adjacent regularization and conversion table regularization can be converted into each other (so long as the child does not use the same surface form for more than one standard form, in which case adjacent regularization would be preferable), and either of them is easily extended to cover all cases covered in CHAT by six widely different methods. Adjacent regularization is more efficient for some purposes.

PRINCIPLES OF SYSTEMATIC DESCRIPTION.

Many of CHAT's above-mentioned troublesome aspects can be traced to three assumptions which derive from a realist rather than relativist perspective on the relation of theory and method. Two of these are reflected in the following quotes (p. 1).

"Transcription is the process of making a written record that corresponds to an audio recording. When transcription is complete, the researcher will not need to make frequent use of the original audio recording." (emphasis mine)

"Coding...is the process of recognizing, analyzing and taking note of phenomena in transcribed speech. It can often be done just on the basis of a reading of a transcript." (emphasis mine)

The first controversial assumption is that the transcript is an objective record of the interaction (cf. "corresponds to"). Critics of that position would note that even the job of preserving words and morphemes cannot be done in a totally

neutral way (as already discussed above, with reference to *wan(t) (t)o*, and *he-'has/he-'is*). Ignoring biases does not make them go away; it only allows them to operate undetected.

The second controversial assumption is that, once it is "complete," the transcript can serve as a substitute for the tape recordings. Critics would note that transcripts are never complete in any absolute sense, but only with reference to a particular perspective at a particular point in time. Transcription systems which have tried to be more exhaustive (Pittenger, Hockett & Danehy 1960) usually figure only minimally in actual analyses, since the enormous extra detail gets in the way of seeing larger patterns in the data. And even these never succeed in preserving every potentially informative aspect of an interaction.

But even if they are complete at some point in time, with respect to *some* particular perspective, many interaction researchers—including for example Emanuel Schegloff (personal communication, 1989), and many others—are adverse to examine any interactions on the basis of a transcript alone, that is, without concurrent access to the actual tape recordings. To do so restricts the sources of insight in the research to only those explicitly encoded in the transcript at a particular point in time, precluding modifications in light of further theoretical developments, as well as the possibility of later correcting inappropriate bias in the transcript. It is the tapes rather than the transcripts which are the primary data.

The primacy of tapes over transcripts has already been demonstrated once in the few decades that tape recording has been possible in child language research. Following the shift in theory and method in the 1970's (discussed above) it would have been desirable to have been able to supplement the original transcripts with more discourse-relevant information (i.e., context, and more detail concerning the spoken language), to see how much and in what ways the added information modified prior conclusions or clarified findings left previously unexplained. Unfortunately, many of those tapes had not survived that long—in some cases, due to physical deterioration of the still new technology (Ervin-Tripp, personal communication, 1986), in other cases (e.g., two of the three children in the Roger Brown project), due to being intentionally over-written with new data gathering, in the then-common belief that once transcribed the tapes were no longer needed (Slobin, personal communication, 1988), making that course of action impossible.

Even when all that is available is the transcript, it is important to be constantly aware of the primacy of tapes over transcripts, to ensure appropriate caution in the conclusions drawn concerning the data, to ask oneself in what ways additional prosodic and contextual information might have modified the results, and to gather additional data where needed. It is only by appreciating the limits of (as well as the value of) archived data that the field can continue to develop as a science of interaction rather than a science of frozen and partial representations of interactions.

The third controversial assumption underlying CHAT is the view that transcription is a cataloguing enterprise, with descriptors being "tags" for self-evident

entities in the world, rather than being constructs imposed on the world by the researcher, whose meaning derives partly from the opposition of each descriptor to the others in the system. Consistent with this, CHAT is basically a juxtaposition of conventions (from diverse earlier projects), rather than a cohesive system in its own right. Particularly important in this respect are its lack of adherence to the following design principles:

- (1) Categories should contrast systematically with one another (orthogonality).
- (2) A single category should not mark unrelated functions (conceptual purity).
- (3) Where more than one category applies, more than one should be marked (multiple marking).

Researchers whose category systems do not satisfy these principles sometimes turn to mathematical or statistical methods to clarify relationships after the fact. But there are three problems with that approach. First, mathematical methods cannot compensate for badly confounded categories. Second, even if they could, it would seem perverse to ignore more systematic oppositions of demonstrated value from prior research. Third, because such categories do not contrast systematically, more of them are required, which clutters the transcript and makes more difficult the work of those researchers rely not so much on mathematical approaches as they do on intuitive insights gained while reading line-by-line through the transcript. This is a time-honored method in the field, and one which has brought much of value. Due to their greater immersion in the data, however, such researchers are also far more affected than others when transcripts are biased inappropriately. Important insights would be missed if we were to turn exclusively to mathematical methods and form-based computer searches in future work. Clearly a transcription standard is needed which can accommodate both types of research.

THE MINIMALIST PROPOSAL.

The recommendations made in this paper constitute the minimalist proposal. They are collectively summarized in the Appendix. This set of distinctions is widespread in the literature though symbols used to mark them may vary. The symbols indicated in the Appendix conform to the conventions found in the (non-CHAT) minimalist data of Eve Clark on CHILDES (Table 6). For purposes of comprehensiveness needed in a field-wide standard, I have also added some of the conventions of Gumperz (1989), DuBois, Cummings, & Schuetze-Coburn (1988), Jefferson (1984) and others, together with explicit conventions concerning regularization.

Insert Table 6 about here

Where these distinctions are found in both discourse research and communication disorders, I have used the symbol found in discourse research. The reason for

this is that most child language research is done with children under five who are not communicatively impaired, with the goal being to describe the child's system in its own terms rather than with reference normative adult categories. Most child language research is therefore ideologically closer to discourse research, so it makes sense that the symbols chosen should also be more consistent with that than with clinical research (see below).

PREREQUISITES FOR COMPUTER ANALYSIS.

In the above discussion, I have repeatedly mentioned the goal of *exhaustiveness* in computer search, that is, retrieving all relevant instances in the data set, and made only passing mention of the other goal in computer search, namely, *economy*, or minimizing false selection or overselection (such as, *only* when searching for *on*). The reason for this difference in emphasis arises from differences in the severity of failing to achieve these two types of goals.

Retrieving extraneous instances, though annoying, is easily detected and ultimately harmless to research conclusions. In contrast, overlooking relevant data threatens the fundamental validity of the research outcome. Without exhaustiveness, what is retrieved can easily be biased samples, and conclusions based on them may mislead more than clarify. In addition, this type of error gives no indication of its presence. Usually it is only detected by chance when reading through the data or scanning results from some other computer search performed on the data.

Until recently computer archives of language data made no provisions for regularization. At present provisions made are only partially adequate in many large archives. Like other systems up to now, CHAT conventions are more oriented toward differentiating items (economy) than toward uniform retrieval of related forms (exhaustiveness). This is apparent in CHAT's sometimes overlapping categories and in the large number of synonymous descriptors, the minimal constraints placed on code syntax, and in the use of separate dependent tiers for things which most researchers would want to retrieve jointly if they were retrieved at all (as discussed above, concerning the three tiers containing contextual information, and the three containing gesture information).

This difficulty is not unique to CHAT conventions but is, rather, widespread, seemingly the result of the relative newness of the methodology of using computers to retrieve utterances in large natural language texts, together with the fact that human beings easily overlook their own talents at regularizing irregular data. In contrast, computers take data very literally. If the computer is told to find instances of *cause*, it will ignore an instance of *cuz*, *cos*, or *'cause* even if a human reader would detect them.

Though its causes are understandable, non-exhaustiveness of search is too dangerous to go unchecked or minimally checked in archive projects. There are basically two ways of minimizing it:

- (1) avoiding unnecessary variability in the data;
- (2) associate a standard form with each nonstandard form (either beside the item or in a conversion table), i.e., "regularization."

"Regularization" has already been treated at length above (under subsections 1 and 6, concerning "Adult morphemes" and "Special form markers"). As for variability, this is another advantage of minimalism. When fewer things are marked, or when they are marked more simply and more systematically, there is less chance for variability, and errors, when they do arise, are more easily detected.

The basic transcription requirements needed for computer assisted analysis of language data are extremely minimal. They include systematic marking (as just discussed), optional selection/suppression of utterances by different speakers, optional selection/suppression of elements within delimiters (e.g., single and/or double parentheses), and optional suppression of such things as differences in upper/lower case letters, punctuation marks, etc.

The alternative proposed in this paper is simpler, more streamlined, and more systematic than CHAT. CHILDES' analysis programs can already accommodate non-CHAT formats of the type proposed here.

CONTINUITY WITH DISCOURSE RESEARCH IN SYMBOL CHOICE.

CHAT authors advocate using different symbols than those already established in discourse literature (as reviewed, for example, in Ehlich & Switalla 1976). In place of .. or ... for pauses (e.g., Labov & Fanshel 1977; Chafe 1980; Gumperz 1982, 1989; Tannen 1984) or + (in second language studies such as the Heidelberger Forschungsprojekt, 1975, and the project described in Wittenburg 1983 and Perdue 1984), CHAT suggests [#], apparently based on the original SALT transcription conventions designed by Miller & Chapman (1983) for normative description and clinical work.²² If, however, most child language researchers prefer child-centered description and approach child language as situated discourse, then it makes more sense to mark pauses in a manner consistent with work in that research tradition, i.e., adult discourse research.

²² In the original version of SALT, brackets enclosed anything which the researcher would want to suppress when computing mean length of utterance (MLU) in morphemes. In recent years, however, morpheme-based MLU counts have fallen into disfavor, especially for languages more heavily inflected than English. Even in English, quantitative criteria for detecting when a particular morpheme becomes productive in the child's speech (rather than being carried along in an unanalysed amalgam with other elements) are no longer as straightforward as they once seemed. Judging something as productive if it occurs in 90% of obligatory contexts (from adult perspective) (cf. Brown 1973), presupposes on the one hand, that the researcher knows what the child meant (singular versus plural; past versus present reference) and on the other, that the adult-centered constraints are also part of the child's system, which may or may not be the case. The current version of SALT is more amenable to discourse research than was the original version, in response to the shifting interests of researchers in the field as a whole.

Because of the excessive number of nonorthogonal (hence inefficient) distinctions required in CHAT, more distinct symbol strings are needed than in other systems. CHAT uses both bracketted and unbracketted versions of some symbols to enlarge the set of symbol strings available for use. This leads to a lack of parallelism among paradigmatic alternatives (e.g., one symbol, ?, for questions and rising intonation, but three symbols, [\], for interruption) as well as similarity of symbols with very different functions (e.g., one symbol, ?, for questions and rising intonation but three symbols, [?], if the researcher is uncertain that the utterance was transcribed correctly). If the distinctions advocated under CHAT were trimmed back to an orthogonal and basic set, the extra square brackets would not be needed, and this would further benefit readability and efficient computer search.

DISCUSSION AND CONCLUSIONS.

In this paper, it is argued that the CHAT transcription system proposed by the CHILDES archive cannot serve as a field-wide standard for child language research because it is biased toward three perspectives no longer current in the field—toward the written rather than spoken language, toward a compositional approach to language and discourse, and toward the view that child language is merely a deficient version of adult language rather than systematic in its own right. In addition, it is argued that CHAT's ad hoc categories, overanalyticity, and non-uniform treatment of regularization hinder efficient and exhaustive computer search even when used with CHILDES' own analysis programs.

This paper proposes in its place a system emerging as a standard in the literature, supplemented by a uniform approach to regularization, which involves preserving both what the child actually said, and the nearest adult equivalent without blurring the distinction between them and without adult-oriented bias.

The goal is to keep categories marked in the transcript to a minimum, chosen to contrast with each other systematically in ways shown by past research to be relevant to child language acquisition, and not included unless likely to be of use to most researchers, since every added distinction takes extra time to make. More specific categories are seen as too problem-specific and theory-committed for a field-wide standard, or added only very cautiously since they prejudice the analysis.

The proposed format conforms to requirements of CHILDES' programs, but has substantive and technical advantages over CHAT, including better readability, less theory-committedness, more systematic and better suited for exhaustive computer search as is needed to avoid unintentionally biased results.

Though particularized to a comparison between CHAT and an alternative, the goal here was to raise issues which transcend any particular system of transcription—issues which are especially timely given the increasing popularity of computer-searchable transcript-based archives, concurrent with an increasing emphasis on prosodic and dynamic aspects of interaction.

REFERENCES

- Atkinson, J. M. & Heritage J. (eds) (1984). *Structures of social action: Studies in conversation analysis*. Cambridge: Cambridge University Press.
- Bloom, L. (1975) *One word at a time*. The Hague: Mouton.
- Bloom, L. (forthcoming). Transcription and coding for child language research. To appear in J. A. Edwards & M. D. Lampert (eds), *Transcription and Coding Methods for Language Research*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Brown, G. & Yule G. (1983). *Discourse analysis*. Cambridge: Cambridge University Press.
- Brown, R. (1973). *A first language: the early stages*. Cambridge, MA: Harvard University Press.
- Chafe, W. (ed) (1980). *The Pear stories: cognitive, cultural, and linguistic aspects of narrative production*. Norwood, N.J.: Ablex.
- Chafe, W. (1987). Cognitive constraints on information flow. In R. S. Tomlin (ed), *Coherence and grounding in discourse*. Philadelphia: John Benjamins Publishing.
- Chafe, W. (1989). Prosodic and functional units of language. To appear in J. A. Edwards & M. D. Lampert (eds) (forthcoming) *Transcription and coding for language researchers*. NY: Erlbaum.
- Cruttenden, A. (1986). *Intonation*. Cambridge: Cambridge University Press.
- Crystal, D. (1975) *The English tone of voice*. London: Edwards Arnold.
- Cutler, A. & Sweeney, D. (1987). Prosody and the development of comprehension. *Journal of Child Language*, 14, 145-167.
- DuBois, J. W., Cumming, S. & Schuetze-Coburn, S. (1988). Discourse Transcription. To appear in J. A. Edwards & M. D. Lampert (eds) (forthcoming) *Transcription and Coding Methods for Language Research*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Edwards, J. A. (in press). Transcription in discourse. In Bright, W. (ed), *Oxford International Encyclopedia of Linguistics*. New York: Oxford University Press.
- Edwards, J. A. & Lampert, M. D. (eds) (forthcoming). *Transcription and Coding Methods for Language Research*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Ehlich, K. & Switalla, B. (1976). Transkriptionssysteme—Eine exemplarische Übersicht. *Studium Linguistik*, 2, 78-105.
- Erickson, F. & Shultz, J. (1982). *The counselor as gatekeeper: Social interaction in interviews*. New York: Academic Press.
- Fletcher, P. (1985) *A child's learning of English*. Oxford: Basil Blackwell.
- Galligan, R. (1987). Intonation with single words: purposive and grammatical use. *Journal of Child Language*, 14, 1-21.

- Gumperz, J. J. (1982). *Discourse strategies*. Cambridge: Cambridge University Press.
- Gumperz, J. J. (1989). Contextualization and understanding. Technical Report of the Cognitive Science Program, University of California, Berkeley.
- Gumperz, J. J. (forthcoming). Transcribing talk in natural settings. In J. A. Edwards & M. D. Lampert (eds) *Transcription and coding for language researchers*. NY: Erlbaum.
- Halliday, M. A. K. (1975). *Learning how to mean: Explorations in the development of language*. London: Edward Arnold.
- Heidelberger Forschungsprojekt "Pidgin-Deutsch" (1975). *Sprache und Kommunikation ausländischer Arbeiter*. Kronberg: Scriptor.
- Jefferson, G. (1984). Transcript notation. In Atkinson, J. M. & Heritage, J. (eds), *Structures of social action: Studies in conversation analysis*. Cambridge: Cambridge University Press, pp. ix - xvi.
- Labov, W. and Fanshel, D. (1977). *Therapeutic discourse: Psychotherapy as conversation*. New York: Academic Press.
- Lea, W. A. Prosodic aids to speech recognition. (1980) In W. A. Lea (ed) *Trends in speech recognition*. Englewood Cliffs, NJ: Prentice-Hall.
- Leech, G. & Svartvik, J. (1975). *A communicative grammar of English*. London: Longman.
- Levinson, S. C. (1983). *Pragmatics*. Cambridge: Cambridge University Press.
- Macaulay, E. K. S. (to appear) "Coz it izny spelt when they say it": displaying dialect in writing. *American Speech*.
- MacWhinney, B. (1988). *CHAT Manual, 16 March 1988: Manual for use of the CHAT Transcription Coding System of the Child Language Data Exchange System*. (unpublished manuscript).
- MacWhinney, B. & Snow, C. (1985) The child language data exchange system. *Journal of Child Language*, 12, 271-296.
- Miller, J. F. & Chapman, R. S. (1983) *Systematic Analysis of Language Transcripts (SALT)* [Computer program]. Madison, WI: University of Wisconsin-Madison, Language Analysis Laboratory, Waisman Center on Mental Retardation and Human Development.
- Ochs, E. (1979). Transcription as theory. In E. Ochs & B. B. Schieffelin (eds), *Developmental Pragmatics*. New York: Academic Press, pp. 43-72.
- Ochs-Keenan, E. (1983) Conversational competence in children. In E. Ochs & B. B. Schieffelin (eds.) *Acquiring conversational competence*. Boston: Routledge & Kegan Paul.
- Perdue, C. (ed) (1984). *Second language acquisition by adult immigrants: a field manual*. Rowley, Mass.: Newbury House Publishers.

- Peters, A. M. (1983). *The units of language acquisition*. Cambridge: Cambridge University Press.
- Pittenger, R. E., Hockett, C. F. & Danehy, J. J. (1960). *The first five minutes*. Ithaca, New York: Paul Martineau.
- Preston, D. (1985). The Li'l Abner Syndrome: Written representations of speech. *American Speech*, 60 (4), 328-336.
- Pye, C. (1988). Refining transcriptions: the significance of transcriber 'errors'. *Journal of Child Language*, 15, 17-37.
- Scollon, R. (1976). *Conversations with a one year old*. Honolulu: The University Press of Hawaii.
- Schwartz, A. H. (1985). Microcomputer-assisted assessment of linguistic and phonological processes. *Topics in Language Disorders*, 6, 26-40.
- Slobin, D. I. & Welsh, C. A. (1973). Elicited imitation as a research tool in developmental psycholinguistics. In C. A. Ferguson & D. I. Slobin (eds), *Studies of child language development*. New York: Holt, Rinehart & Winston, pp. 485-497.
- Svartvik, J. & Quirk, R. (eds) (1980). *A corpus of English conversation*. Lund: CWK Gleerup.
- Tannen, D. (1984). *Conversational style*. Norwood, N.J.: Ablex.
- Tannen, D. (1981). Review of: Therapeutic discourse: Psychotherapy as conversation. by W. Labov and D. Fanshel. *Language*, 57 (2), 481-6.
- Tedlock, D. (1983). *The spoken word and the work of interpretation*. Philadelphia: University of Pennsylvania Press.
- Waibel, A. (1988). *Prosody and speech recognition*. San Mateo, California: Morgan Kaufmann Publishers.
- Wittenburg, P. (1983). ESF Project—Ecology of Adult Language Acquisition (EALA)—Data Processing Aspects. (Unpublished; instructions to project members.) Max-Planck-Institut für Psycholinguistik, Nijmegen, April 1983.

Table 1
CHAT's Utterance Delimiters,
 copied verbatim from MacWhinney (1988: 11-12)

4.0.2 Utterance Delimiters

A variety of special symbols are permitted on the main line. The syntax and use of these special symbols is tightly defined. The first set we will discuss are the delimiters. A delimiter simply marks the end [sic] of an utterance. The CHAT coding scheme does not attempt to tell the transcriber how to decide what should constitute an utterance. However, it does require that the text be segmented into utterances. Each utterance on the main line must end with an utterance delimiter.

Period .

A period marks the end of an unmarked (declarative) utterance. When you are unsure which delimiter to place after an utterance, use a period. If none of the previous four [sic] delimiters seem to fit, use a period. Omit periods when indicating an abbreviation. Here are some examples of unmarked utterances.

*SAR: I got cold. pickle. no.

Question ?

The question mark indicates the end of a question. A question is an utterance which uses a wh-question word, subject-verb inversion, a tag or question intonation (in English, rising intonation). It serves as a request for information or action. Here is an example of a question.

*FAT: is that a carrot?

The question mark can also be used after a declarative sentence when it is spoken with the rising intonation of a question.

Exclamation !

An exclamation point marks the end of an utterance which carries emphatic sentential stress. Often imperatives are marked with an exclamation point. Here is an example of an exclamation.

*MOT: sit down!

Trailing off [...]

The trailing off marker is the delimiter for an incomplete but not interrupted utterance. When speakers trail off, they leave their utterances incomplete. Their interest may have shifted from what they are saying to something else or

they might have forgotten what they were going to say. In the case of trailing off, speakers are NOT interrupted by another speaker or even by themselves. This is an example of a trailed off utterance.

*SAR: smells good enough for [...]

*SAL: what is that?

Interruption [N]

This symbol is used for an utterance which is incomplete because another speaker begins talking or some event occurs causing the speaker to break off. One speaker is interrupted by another speaker. Here is an example of an interruption.

*MOT: what did you [N]

*SAR: Mommy.

The notion of interruption is to be taken very broadly to include both invited and uninvited interruptions.

Latching [>]

The delimiters we have discussed so far mark the ends of utterances. Other delimiters can be used to mark special aspects of the beginning of utterances. One of these is "latching." Latching occurs when an utterance of one speaker follows quickly on the heels of the last utterance of the preceding speaker without the customary short pause between utterances. An example of this is:

*MOT: why did you go?

*SAR: [>] I really didn't.

Completion [+]

The symbol [+] marks the continuation and completion of an incomplete utterance either by Speaker A after interruption by Speaker B, as in:

*CHI: so after the tower [N]

*EXP: yeah.

*CHI: [+] I go straight ahead.

Or it can be used to mark the completion of the utterance of Speaker A by Speaker B, as in:

*HEL: if Bill had known [N]

*WIN: [+] he would have come.

Table 2
Example of a Prosodically-based Alternative
to CHAT's Unit-final Delimiters
(from DuBois, Cummings & Schuetse-Coburn 1988)

Distinctions marked	Symbols Used	
Rising intonation (regardless of syntactic structure)	?	Question mark
Intonation suggesting finality (in English, falling)	.	Period
Non-rising/non-finality, sometimes called "steady" intonation	,	Comma
Non-completed word	-	Tilde
Non-complete intonation contour	--	Double-tilde
Overlapping speech by multiple speakers	[]	Enclosed in brackets

Table 3

A comparison between Lois Bloom's (1975:206) original format (on the left), and CHILDES' version of those data in CHAT format (on the right).

<i>Allison III: 20 months, 3 weeks</i>			
1. (M and A sitting on chair; A wearing half-zippered jacket; fingers in her mouth)		@Begin	
What did you see? What did you see over there?		@Participants: ALI Allison child, MOT mother	
(M points to monitor)		@Age of ALI: 20 months, 3 weeks	
(A looking at monitor with fingers in her mouth)	Mommy/	@Sex of ALI: Female	
Yeah. And who else?		@Filename: allison3.cha	
Baby doll?	baby doll/	*MOT: what did you see? what did you see over there?	
Baby Allison. Yes. What's Allison doing?	baby Allison/	%sit: Mother and Allison sitting on chair; Allison wearing half-zippered jacket; fingers in her mouth	
What are you doing?		%gpx: <aft> points to monitor	
(A's fingers are in her mouth and she opens and closes her mouth as she watches the monitor)		*ALI: Mommy.	
	play/	%gpx: looking at monitor with fingers in her mouth	
Play? I don't see you playing. I don't see you playing.		*MOT: yeah. and who else?	
	talk/ talk/ (whisper)	*ALI: baby doll.	
Oh, talk. Yes, yes.		*MOT: baby doll?	
		*ALI: baby Allison.	
2. Are you ready to take your coat off?	no/	*MOT: baby Allison. yes. what-'is Allison doing? what are you doing?	
No? Let's unzipper it, it's very warm.		*ALL: play.	
(M tries to unzip A's coat; A resists)	no! on/	%act: <.5u> Allison's fingers are in her mouth and she opens and closes her mouth as she watches the monitor	
		*MOT: play? I do-'nt see you playing. I do-'nt see you playing.	
		*ALI: talk. talk.	
		%par: whisper	
		*MOT: oh, talk. yes, yes. are you ready to take your coat off?	
		*ALI: no.	
		*MOT: no? let-'us unsipper it, it-'is very warm.	
		%act: <aft> tries to unsip Allison's coat	
		*ALI: no! on.	
		%act: resists	

Table 4

Lois Bloom's data modified only as necessary for line-oriented search,
i.e., no added interpretive categories not contained in the original.

@Begin

@Participants: A = Allison (child), M = Mother

@Age of A: 20 months, 3 weeks

@Sex of A: Female

@Filename: allison3

%(M and A sitting on chair; A wearing half-zippered jacket;
%fingers in her mouth)

M: what did you see? what did you see over there?

%(M points to monitor)

%(A looking at monitor with fingers in her mouth)

A: Mommy/

M: yeah. and who else?

A: baby doll/

M: baby doll?

A: baby Allison/

M: baby Allison. yes. what's Allison doing? what are you doing?

%(A's fingers are in her mouth and she opens and closes her
%mouth as she watches the monitor)

A: play/

M: play? I don't see you playing. I don't see you playing.

A: talk/ talk/ (whisper)

M: oh, talk. yes, yes.

are you ready to take your coat off?

A: no/

M: no? let's unzipper it, it's very warm.

%(M tries to unzip A's coat; A resists)

A: no!/ on/

Table 5
CHAT's "Special Form" tags, definitions, category overlap,
and occurrences of each tag found in CHILDES' data files

CHAT Tag (and overlapping categories)	Uses of the tag in CHILDES	Uses of other tags in CHILDES	Meaning (as needed)
<i>ai</i> letter	ai! bi! ci! 's		A. B. C.'s
<i>ab</i> babbling (see <i>am</i> , <i>aw</i> , <i>ap</i>)	goo+gooab, ga+gaab	gooooom, gagaaw, gagaap	
<i>am</i> motherese (see <i>ab</i> , <i>aw</i> , <i>ac</i> , <i>af</i> , <i>as</i>)	choo-chooam, choochooam = train jamaam, beddy-byoam Grampaam, grammaam bookoam, blankoam	choo+chooam jamaam, bed+byeac Grampicaf, Grampaam	pajamas, bedtime book, blanket
<i>aw</i> wordplay (see <i>@b</i>)	tralalaaw, whoo/duddooaw, gagaaw	gagaap, ga+gaab	
<i>ao</i> onomatopoeia (see <i>am</i> , <i>aw</i> , <i>ac</i>)	kcrplonkao, mooorio, splashao, bangao	bangaac	
<i>ah</i> hesitation (see <i>as</i> , <i>aq</i>)	umah, umah	umaa, umaq	
<i>aq</i> qualifier (see <i>ab</i> , <i>am</i> , <i>ac</i> , <i>as</i>)	hahaha/laughaq, hahaha/laughingaq mhm-q-yes, hm-q-ques uhohaq-trouble, ahoh/troubleaq yeah/yes+airaq, yeah/yeahq mnyumaq, yum/deliciousaq, mnytastyaq cm/themaq, umaq	yum+yumam, yumac cm/themaas, umaa, umah	them, hesitation
<i>an</i> neologism = "derived from existing forms, including over-regularization"; in Sachs' data also some wordplay (see <i>ac</i> , <i>ae</i>)	fromanien, poisonien, slipperyien comedien, hidodien, flyodien, sittedien tearedien, sneakedien supercalifragilisticexpialidociousien	comediac, hidodiac, flyodiac, sittediac tearediac, sneakediac	came, hid, flew, sat tore, sneaked
<i>ad</i> derivative from adult (see <i>ae</i>)	winahadad, slippyad, datad	datac	windshade, slippery, that
<i>ae</i> child invented forms = "not derived from other words by standard derivational processes" (see <i>am</i> , <i>aw</i> , <i>ao</i>)	taperac, weigherac, weighterac peasyac, tumblesaultac, buffryac glaciac, corn-butterac, bangaac mooshyac, wawaac, homiac	bangao gocahyam, wa+waam, horsicam	tapererecorder, scale pacifier, sommersault, butterfly glasses, butter for corn a drink, horse
<i>ae</i> extended English (see <i>aq</i> , <i>am</i> , <i>ac</i> , <i>ad</i>)	bonkac, yuckyac, nogginac, grandcheitac thingyac, denac, datac	yukky/badaq thingyam, thingyac, denac, datad	thing, then, that
<i>ap</i> phonological consistent form (see <i>ac</i> , <i>ae</i> , <i>ab</i>)	agaap, gagaap squishedap	gagaab squashedac, smushedac, squocahac	pointing morph, babbling
<i>af</i> familial forms (see <i>am</i> and <i>as</i>)	Grampicaf, Gammicaf, fuzzyaf	Grampaam, Grammaas	grandparents, a drink
<i>as</i> abbreviation (see <i>as</i> , <i>af</i> , <i>aq</i> , <i>ac</i> , <i>am</i> , <i>af</i>)	causaa 'causaa, tillaa, tillaa, (un)tillaa anvaa, cm/themaas, umaa haftaa, gonnaas, gunnaas Grammaas, Grandpaas	anah, cm/themaq, umah, umaq havtaaf, havaac Gammicaf, Grampaam	cause, until and, then, hesitation have to, going to grandparents
<i>at</i> second language form	istenermas, unowas, dcaas		my God, one, two

Table 6

Sample from Eve Clark's data on CHILDES,
in a non-CHAT minimalist format, related to that proposed here.

Age 2;4,25

Date June 30, 1976

-----SESSION NINE

S = Shem (child)

C = Cindy (observer)

M = Mother

C: let's press these two down and keep 'em down, okay? m'kay?

%pressing taperecorder buttons

S: xxxxxx.

C: no, you can't push that one when these are moving.

ooh, what do you got on your hands here?

were you playing with something?

S: i got a bandaid.

C: oh, what did you get a bandaid from?

S: i goi(ng) \$ pri- \$ prick- my kitty pricked my... hurt.

C: he pricked you and it hurt?

S: \$...\$...\$...an' i got a bandaid.

C: what was it for?

S: to the hurt.

C: oh, did you go to the doctor?

S: no, i go in my house- i go in my backyard- go in \$ grass an' \$ hurts.

C: oh, in the grass. oh, what do you think...

S: ...my kitty.

C: with your kitty?

S: yeah.

C: she did it?

S: yeah.

C: did your kitty do it?

S: yeah.

C: what did she do to you?

Appendix

Minimalist Alternative to CHAT¹

1. **Background information**—indicated in file headers or 00readme file.
2. **Comments, and nonverbal and contextual information** is indicated either in double parentheses on the same line (if brief and pertinent to only that utterance):

A: Where's the doggie? ((whispering))

otherwise, placed on a separate line beginning with a percent sign:

%S picks up the puzzle piece

3. Utterances are preceded by speaker identifiers (one to three initials, followed by colon; more speaker information given at top of file).

Utterances are transcribed in standard spelling, except for:

- Inaudible words, transcribed as xxx, or one x per syllable.
- Unclear words, (bird), i.e., best guess in parentheses
- Untranscribed segments ((untrans. 2 sec.)), with time estimates
- Spoken forms which lack standard written conventions (e.g., *gonna*) or
- Lexical or syntactic departures from the adult system (e.g., *breaked*)

In last two cases, preserve the form the child used, and associate it with a standard form for search purposes, either adjacent to it, set off in parentheses with an equal sign:

He's gonna (=going to) bark.

or as part of an exhaustive conversion table in a "lexicon" file:

gonna, goin', gon' tuh = going to

cause, 'cause, cuz, cos = because

tuh, ta = to

Append the following minimal markings to the end of the item for search purposes:

- * Lexical/syntactic departures from adult form:

I hitted* (=hit) it.

I been* sit* (=sitting) all day.

- "Omitted" form: I want " (=to) go.

- | Lexical innovation: I want corn-butter: (=butter for corn)

4. **Basic Prosodics.** Utterances are broken into "intonation" units, i.e., spurts of speech, falling under a coherent intonation contour and often set off by a pause. For each intonation unit, prosodic contour is indicated by punctuation marks:

¹ Individual categories and symbols used here are drawn from systems used by E. Clark in her data on CHILDES, Gumpers (1989), Chafe (1987, 1989), DuBois, Cummings, & Schuetse-Coburn (1988), Jefferson (1984), and others cited in the text.

. Terminal fall (in English, falling)
 ? Interrogative rise (regardless of syntactic structure)
 , Continuing or listing contour (slight pause, with sustained level tone, or slight falling but non-terminal tone)
 ! Emphatic tone
 " Word or syllable stress

.. Pause, if barely perceptible
 ... Pause of approximately half a second (0.3 - 0.6 seconds)
 <n> Pause longer than 0.7 seconds (n = seconds)
 - Truncation (e.g., what ti- what time is it?)
 [1] Overlap, for example:

S: I see a [doggie!1]
 C: [Yes, dear.1]

== Latching, for example:

R: so you understand the requirements?
 B: ==yeah, i understand them.

5. Provisions for exhaustive search.

- Regularization of all nonstandard forms, i.e., associating a standard form with every nonstandard variant in either of the ways described under 1. above.
- Avoiding unnecessary variability in the data file, due to any of the following: inconsistent syntax, misspelling, inconsistent use of spaces, tabs, and control characters, formatting commands, inconsistent capitalization, hyphenation.
- Do not add codes you don't think you will use in your research. For codes you do add: if categories are not mutually exclusive, then allow multiple coding; if categories are mutually exclusive, define explicitly with examples of prototypical instances as well as of exceptional cases. Check consistency of code use repeatedly.